

University of Groningen

Domain structure and pore loops in the 2-hydroxycarboxylate transporter family

Lolkema, Juke S.

Published in:
Journal of Molecular Microbiology and Biotechnology

DOI:
[10.1159/000095634](https://doi.org/10.1159/000095634)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2006

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Lolkema, J. S. (2006). Domain structure and pore loops in the 2-hydroxycarboxylate transporter family. *Journal of Molecular Microbiology and Biotechnology*, 11(6), 318-325. <https://doi.org/10.1159/000095634>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Domain Structure and Pore Loops in the 2-Hydroxycarboxylate Transporter Family

Juke S. Lolkema

Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

Key Words

Secondary transport • Membrane protein • Domain structure • Pore loop • Structural classification • MemGen • Membrane topology

Abstract

The 2-hydroxycarboxylate transporter (2HCT) family is a family of bacterial secondary transporters for substrates like citrate, malate and lactate. The family is in class ST[3] of the MemGen classification system that groups membrane proteins in structural classes based on hydropathy profile analysis. The combination of computational analysis of the proteins in class ST[3] and available experimental data on members of the 2HCT family has yielded a detailed structural model of the transporters. The core of the model is formed by two homologous domains with opposite orientation in the membrane. Each domain consists of 5 *trans* membrane segments and contains a pore loop between the 4th and 5th segment. The two pore loops enter the membrane-embedded part from opposite sides of the membrane (*trans* pore loops) and are believed to form the translocation pathway in the 3D structure. A genome wide study of the cellular location of the C-terminus of all *Escherichia coli* membrane proteins [Daley et al., 2005. Science 308: 1321–1323] showed that the C-termini of the 19 *E. coli* proteins in class ST[3] were correctly predicted by the structural model.

Copyright © 2006 S. Karger AG, Basel

Introduction

The 2-hydroxycarboxylate transporter (2HCT) family of secondary transporters contains 54 unique members that are all found in the bacterial kingdom. The characterized members of the family are transporters for citrate, malate and lactate, substrates that contain the 2-hydroxycarboxylate motif, hence the name of the family. Like in other families of secondary transporters, the members of the 2HCT family represent different modes of energy coupling. The transporters are either H⁺ or Na⁺ symporters or they catalyze exchange between two substrates. Well-studied members of the family are listed in table 1 (for a recent review, see Lolkema and Sobczak [2005]).

The transporters of the 2HCT family function in citrate and malate breakdown pathways, mostly under anaerobic conditions. The pathways are characterized by decarboxylation steps in which oxalate is converted to pyruvate and malate to pyruvate or lactate in the citrate and malate breakdown routes, respectively. Na⁺-coupled citrate transporters like CitS of *Klebsiella pneumoniae* and CitC of *Salmonella enterica* found in the γ subdivision of the phylum Proteobacteria are involved in the fermentative degradation of citrate to acetate and carbon dioxide yielding ATP. H⁺-coupled uptake of malate by MalP of *Streptococcus bovis* in the phylum Firmicutes results initially in the formation of pyruvate and carbon dioxide. The exchangers MleP and CitP are involved in

Table 1. Characterized members of the 2HCT family

Transporter	Bacterium	Substrates	Transport mode	Function	Reference
CitS	<i>Klebsiella pneumoniae</i>	citrate	Na ⁺ symport	citrate fermentation	Dimroth and Thomer, 1986
CitC	<i>Salmonella typhimurium</i>	citrate	Na ⁺ symport	citrate fermentation	Ishiguro et al., 1992
CitW	<i>Klebsiella pneumoniae</i>	citrate, acetate	exchange	citrate fermentation	Kästner et al., 2002
MleP	<i>Lactococcus lactis</i>	malate, lactate	exchange	malolactic fermentation	Molenaar et al., 1993
CitP	<i>Leuconostoc mesenteroides</i>	citrate, lactate	exchange	citrolactic fermentation	Marty-Teyssset et al., 1995
CimH	<i>Bacillus subtilis</i>	citrate, malate	H ⁺ symport	unknown	Krom et al., 2001
MalP	<i>Streptococcus bovis</i>	malate	H ⁺ symport	malate fermentation	Kawai et al., 1997
MaeN	<i>Bacillus subtilis</i>	malate	Na ⁺ symport	growth on malate	Wei et al., 2000

unique metabolic energy-generating pathways that are found in lactic acid bacteria in the phylum Firmicutes. The transporters are special in that they generate membrane potential during turnover because of a charge difference between the two substrates. MleP couples the uptake of divalent malate to the excretion of monovalent lactate. Lactate is produced inside the cell from the decarboxylation of malate by malolactic enzyme. The decarboxylation reaction consumes a cytoplasmic proton which results in alkalization of the cytoplasm. Overall, the pathway that consists of only the exchanger and the decarboxylase generates proton motive force in the form of both a membrane potential and a pH gradient across the cytoplasmic membrane. Similarly, CitP generates membrane potential by exchanging divalent citrate for monovalent lactate. The pathway leading to the end product lactate is more complex, involving more enzymes and interaction with the glycolytic pathway. Lactate is formed by reduction of pyruvate, the direct product of citrate metabolism in lactic acid bacteria, using reduction equivalents produced in glycolysis. Citrate is cleaved by citrate lyase yielding acetate and oxaloacetate, which is decarboxylated yielding pyruvate. The latter step results in the transmembrane pH gradient. The malolactic and citrolactic fermentation pathways as they are called function as indirect proton pumps. The pathways have important biotechnological implications in the dairy industries and in wine-making.

Secondary transporters are widely distributed in nature and they come in a great genetic and structural diversity, probably reflecting many different translocation mechanisms [Sobczak and Lolkema, 2005b]. The many transporter families listed in the Transporter Classification (TC) system developed by the Saier laboratory [Saier, 2000] require further classification to identify distant relatives with similar folds and mechanisms. In the TC sys-

tem, families are grouped in superfamilies based on sequence analysis using randomizing techniques. We have classified families of membrane proteins in structural classes using hydropathy profile analysis [Lolkema and Slotboom, 1998, 2003]. Here we discuss the relation of the 2HCT family to other families of secondary transporters in the MemGen structural classification system. Together with a large body of experimental data, the analysis results in a detailed structural model for the transporters in the 2HCT family.

MemGen Membrane Protein Classification

Hydropathy profiles of membrane proteins have been used extensively to predict the secondary structure of membrane proteins. The alternating regions of the polypeptide chain traversing the hydrophobic membrane and those residing in the hydrophilic water phase result in typical peak patterns in the hydropathy profile of the amino acid sequence that, in principal, allows for a simple prediction of the topology of the protein in the membrane. The MemGen method takes the information content of hydropathy profiles one step further by assuming that a particular three-dimensional fold of a membrane protein produces a characteristic overall hydropathy profile [Lolkema and Slotboom, 1998]. Then, similar hydropathy profiles group proteins with similar structures which may be the result of divergent evolution of the proteins in the group. The advantage over sequence analysis is that the structure of proteins is better conserved during evolution than the amino acid sequence of the proteins. It follows that, in principle, hydropathy profile analysis is able to detect more distant evolutionary relationships between proteins than sequence analysis.

The starting point for the MemGen classification scheme are membrane protein families that contain homologous proteins based on sequence similarity. Two levels corresponding to evolutionary distance are discriminated. Subfamilies contain sequences that show significant overall sequence identity in a multiple sequence alignment, and families that contain one or more subfamilies. Members in different subfamilies in a family show significant sequence identity in a local sequence alignment as produced by the BLAST algorithm [Altschul et al., 1997]. For more detailed criteria to discriminate families and subfamilies, see Lolkema and Slotboom [2003]. The multiple sequence alignments are used to produce averaged hydropathy profiles of the (sub)families (the 'family profile'), and software has been developed to compare profiles of different families by computing the optimal alignment of the profiles [Lolkema and Slotboom, 1998]. A criterion for structural similarity is obtained by comparing the difference between the aligned profiles with the average difference of the individual member profiles in the two families with the corresponding family profile. It is argued that if the difference between two family profiles is in the same order as the difference between the member profiles and the average profiles, the family profiles might as well be members of each other's families. Families with similar hydropathy profiles are grouped in structural classes that relate membrane proteins that by sequence identity would be non-homologous.

Four structural classes of secondary transporters were identified and described in some detail using the MemGen method: ST[1], ST[2], ST[3], ST[4] in which ST is short for Secondary Transporter [Lolkema and Slotboom, 1998, 2003, 2005]. ST[1] corresponds largely to the Major Facilitator Superfamily (MFS) [Saier et al., 1999] in the TC classification system. ST[2] contains among others the APC superfamily [Jack et al., 2000], the Na⁺- and Cl⁻-dependent neurotransmitter transporters and the Na⁺ glucose transporter family. ST[3] contains the anion transporter (IT) superfamily [Prakash et al., 2003], and, finally, ST[4] contains the glutamate transporter family [Slotboom et al., 1999a, b]. Recent crystal structures of secondary transporters have confirmed distinct structures for structural classes ST[1], ST[2] and ST[4] [Abramson et al., 2003; Huang et al., 2003; Yamashita et al., 2005; Yernool et al., 2004]. The goal of the MemGen classification is to extract all proteins that belong to one of the classes from the NCBI nr protein database which has been accomplished for structural classes ST[3] and ST[4] [Lolkema and Slotboom, 2003, 2005]. ST[3] groups

67 subfamilies in 35 families and contains 3,538 unique proteins. ST[4] is much smaller containing 7 subfamilies in 2 families and 840 unique proteins.

The MemGen classification is not a secondary structure prediction method per se. The method merely identifies families with a similar global fold without giving any details on that fold. However, as a consequence, structural information on the proteins in one of the families bears on all the proteins in the other families in the same structural class. The structural information may range from a crystal structure to a membrane topology model. The transporters of the 2HCT family in structural class ST[3] have been studied in detail resulting in a structural model and a proposal for a translocation mechanism (reviewed in Sobczak and Lolkema [2005a]). Using the optimal alignments of the family hydropathy profiles as a guide, the structural elements of the 2HCT model may be projected on the 35 other families in structural class ST[3] yielding structural models for these families together with a translocation mechanism.

Structural Model of 2HCT Transporters

The transporters in the 2HCT family are integral membrane proteins consisting of about 440 amino acid residues. The structural model for the proteins that is depicted in figure 1 represents a new fold that contains structural elements that are observed in known structures of other integral membrane proteins. The core of the structure is formed by two homologous domains that are connected by a large hydrophilic loop that resides in the cytoplasm. Sequence identity between the two domains is below the level of detection. The domains contain 5 transmembrane segments (TMSs) each and they have opposite orientations in the membrane. Homologous domains with inverted topologies are becoming a common structural motif in membrane proteins and are also observed in the crystal structures of aquaporin [Fu et al., 2000; Murata et al., 2000], the ammonium transporter AmtB [Khademi et al., 2004], the Na⁺/H⁺ antiporter NhaA [Hunte et al., 2005] and the leucine transporter LeuT [Yamashita et al., 2005]. They are likely to originate from a duplication of an internal gene fragment coding for an odd number of TMSs. The same structural motif may also be formed in multimeric proteins, either as homodimers or heterodimers in which the two subunits are homologous proteins [Rapp et al., 2006; Bowie, 2006; Pornillos and Chang, 2006]. With separate subunits, the number of TMSs in each subunit may be odd

or even. An example of the latter would be the drug transporter EmrE that forms a homodimer with inverted topology with 4 TMSs per subunit [Pornillos et al., 2005]. In the structural model of the transporters in the 2HCT family, the loops between the 4th and 5th TMSs in each domain fold back in between the TMSs and form so-called re-entrant or pore loops. The pore loop in the N-terminal domain (region VB) enters the membrane-embedded part from the periplasmic side of the membrane, the one in the C-terminal domain (region XA) from the cytoplasmic side (*trans* pore loops). The two re-entrant loops are believed to be in close vicinity in the 3D structure and to form the translocation pathway for co-ions and substrates. The binding site is believed to be positioned at the membrane-cytoplasm interface where an arginine residue interacts directly with the bound substrate. Pore loops are well-known structures in channel proteins but are also observed in the crystal structure of a glutamate transporter homologue of *Pyrococcus horikoshii* [Yernool et al., 2004]. The two domains form the core of the structure of the transporters in the different families of ST[3] and are likely to be the minimal catalytic unit. Different families may have additional TMSs at the N- or C-termini or in between the two domains. The transporters of the 2HCT family have one additional TMS at the N-terminus locating the latter in the cytoplasm.

***E. coli* Membrane Proteome: Membrane Topology Prediction of ST[3] Proteins**

Between 20 and 25% of the genes in the *E. coli* genome code for membrane proteins. In a major effort, the group of von Heijne [1999] determined experimentally the cellular disposition of the C-terminus of a significant part of the *E. coli* integral membrane proteins [Daley et al., 2005]. In doing so, they generated an invaluable experimental dataset for the validation of membrane topology prediction methods. The cellular location of the C-termini was determined using the reporter gene fusion technique. Alkaline phosphatase (PhoA fusions) and green fluorescent protein (GFP fusions) were fused to the C-termini as positive reporters of a periplasmic and cytoplasmic location, respectively. By using the double reporter technique, the C-terminal location of 502 of the 665 cloned proteins could be determined with high fidelity.

TMHMM is widely used for secondary structure prediction of membrane proteins [Krogh et al., 2001]. The

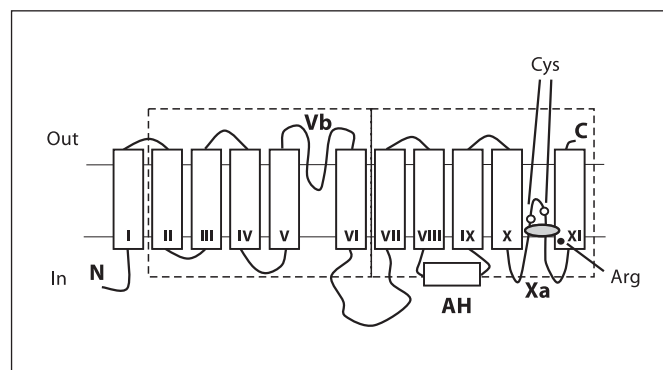


Fig. 1. Structural model for the transporters of the 2HCT family. Boxes represent transmembrane segments. The two homologous domains were indicated in dashed boxes. VB and XA indicate the pore-loop structures and AH an amphipathic surface helix. The putative substrate-binding site in between TMSs X and XI (oval), the conserved arginine residue in TMS XI (●) and the two cysteine residues in pore-loop XA of CitS of *K. pneumoniae* (○) were indicated.

method is based on a Hidden Markov Model and performs well relative to other predictors [Cuthbertson et al., 2005]. Accordingly, TMHMM predicted the cellular location of the C-terminus of 78% of the *E. coli* membrane proteins in the experimental set correctly (fig. 2). Structural class ST[3] in the MemGen classification contains 40 transporters from *E. coli*, 26 of which are found in the Colibri database that was the starting point for the selection of the inner membrane proteome of *E. coli*. The remaining 14 transporters are either found exclusively in different strains or are plasmid encoded. A C-terminal location was obtained for 19 of the 26 proteins (table 2). The 19 proteins cover 7 different families in class ST[3], the 2HCT family not being one of them. Remarkably, TMHMM scores very poorly for this subset of proteins. The C-terminus was predicted correctly for only 8 of the 19 proteins (42%), which is as good (or actually as bad) as flipping a coin. In contrast, the prediction made by the MemGen classification using the structural model of the transporters in the 2HCT family as the basis, predicts the C-terminal location of this set 100% correctly (fig. 2a). The proteins in the subset are all in the ~20% of the membrane proteins that have the C-terminus in the periplasm, which is a characteristic feature of the ST[3] transporters (fig. 1).

The authors of the *E. coli* inner membrane proteomic study promote the idea that prediction methods may improve considerably when a site in the protein (i.e., the C-terminus) is locked in the periplasm or cytoplasm. This

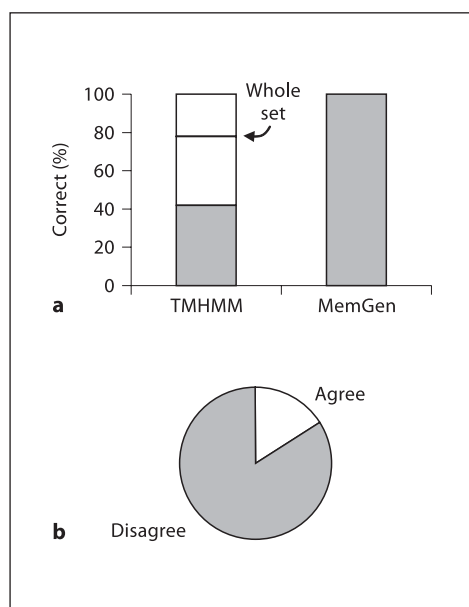


Fig. 2. Secondary structure prediction by TMHMM and MemGen. **a** Percentage of correctly predicted C-terminal locations by TMHMM and MemGen for a subset of 19 *E. coli* membrane proteins found in structural class ST[3] of the MemGen classification (shaded). The experimental data was taken from Daley et al. [2005]. The score for TMHMM on the whole set of *E. coli* membrane proteins was also indicated. **b** Pie graph showing the comparison of the overall membrane topology of the 19 proteins predicted by MemGen and by TMHMM with constraint C-termini (<http://www.sbc.su.se/~melen/TMHMMfix/>).

will be the case when the algorithm has difficulties to decide upon the orientation of the protein in the membrane as a whole because of a 'weak' positive-inside rule. In the subset of 19 proteins, constraining the C-terminus in the periplasm in TMHMM, resulted for only 3 proteins in the same membrane topology model as predicted by MemGen (fig. 2b). Clearly, both methods produce basically different topology models with different numbers of TMSs. TMHMM does not recognize pore-loop structures, and the corresponding regions in the sequences are often interpreted as being transmembrane (see also Cuthbertson et al. [2005]). It follows that while the overall performance of TMHMM is quite good, it fails for certain classes of membrane proteins. For one of these classes, the ST[3] proteins, the prediction based on the MemGen protein classification results in the correct localization of the C-termini for the ST[3] proteins. Identifying such classes and excluding them from the analysis by TMHMM will significantly improve the fidelity of the prediction for the remainder of the membrane proteome.

Pore-Loop Structures in the 2HCT Transporters

Experimental evidence for the pore-loop structures in the structural model of the transporters in the 2HCT family comes from studies of loop XA in between TMSs X and XI (fig. 1) in the Na^+ -citrate transporter CitS of *K. pneumoniae* and the H^+ -citrate/malate transporter CimH of *Bacillus subtilis*. In the absence of a 3D crystal structure, pore loops in membrane proteins are detected by accessibility studies of cysteine residues by thiol reagents with different physicochemical properties. The reagents differ mainly in molecular size and in membrane permeability. The former allow for the detection of cysteine residues in space-restricted but water-accessible cavities of the protein, while the latter allow for the accessibility of a residue from one particular side of the membrane. Clearly, as pore-loop structures fold back in between the TMSs, their accessibility from the water phase is likely to be restricted, and, since they are believed to be part of the translocation site, they may be accessible from both sides of the membrane in different states of the catalytic cycle. Using this approach, two pore-loop structures were identified in transporters of the glutamate transporter family, findings that were later confirmed by the crystal structure of one of the members [Grunewald et al., 1998; Slotboom et al., 1999a, b; Grunewald and Kanner, 2000; Yernool et al., 2004].

Loop XA in CitS of *K. pneumoniae* contains two cysteine residues, the accessibility of which was investigated in single Cys constructs of the transporter [Sobczak and Lolkema, 2003, 2004]. A bulky, membrane-impermeable thiol reagent could only access the thiols from the cytoplasmic side of the membrane, while small membrane-impermeable thiol reagents with sizes similar to the citrate molecule could access the thiols from both sides of the membrane. The experiments demonstrated that an access pathway exists for small molecules from the periplasmic side of the membrane to the cysteine residues in the putative cytoplasmic loop. Similar observations were made for two engineered cysteine residues in loop XA of CimH of *B. subtilis* [Krom and Lolkema, 2003]. Functional importance of the pore loop in the XA region followed first of all from the inactivation of the transporters upon modifying the cysteine residues. Additional data linked the accessibility of the cysteine residues directly to the catalytic state of the transporters. Binding of substrate to the transporter affected the accessibility of the cysteine residues differently when attacked from the periplasmic or cytoplasmic side of the membrane. In the case of CitS, where the transporter can be locked in the binary

Table 2. Experimental topological determinants of ST[3] proteins

MemGen/ST[3] ^a			Colibri ^b	C-terminus			Topology	
family	subfamily	protein		experiment ^c	TMHMM ^d	MemGen	TMHMM _{fix} ^e	MemGen
[st302]ArsB	ArsB1	ARSBecol	arsB	out	in	out	7/5 ^f	5/5 ^f
	ArsB3	YBIREcol	ybiR	out	out	out	4/5	4/5
[st303]AIT	AIT1	YGJEcol	ygjE	out	in	out	7/6	6/5
	AIT1	YBHIecol	ybhI	out	out	out	8/7	6/5
	AIT3	YFBSecol	yfbS	out	in	out	6/4	5/5
[st304]GNT	GNT1	GNTPecol	gntP	out	in	out	5/5	5/5
	GNT1	YGBNecol	ygbN	out	in	out	5/5	5/5
	GNT1	IDNTecol	idnT	out	in	out	5/6	5/5
	GNT1	YJHFecol	yjhF	out	in	out	6/5	5/5
	GNT1	CAD42064ecol	dsdX	out	out	out	6/6	5/5
	GNT1	GNTUecol	gntU	out	in	out	6/5	5/5
	GNT1	AAA58213ecol	gntT	out	in	out	6/4	5/5
	GNT1	AAA23826ecol	dcuB	out	out	out	5/6	5/5
[st305]DCUA	DCUA1	AAA23826ecol	dcuB	out	out	out	5/6	5/5
[st309]DCUC	DCUC1	DCUCecol	dcuC	out	out	out	6/6	5/5
		DCUDecol	yhcL	out	in	out	6/6	5/5
[st310]ATO	ATO1	ATOEcol	atoE	out	in	out	5/5	4/5
[st324]GLTS	GLTS1	GLTSecol	gltS	out	in	out	3/5	5/5
[st325]2HMCT	2HMCT1	YGHKecol	yghK	out	out	out	7/6	5/2/5
		LLDPecol	lldP	out	out	out	7/5	5/2/5

^a <http://molmic35.biol.rug.nl/>^b <http://genolist.pasteur.fr/Colibri/>^c Daley et al., 2005.^d <http://www.cbs.dtu.dk/services/TMHMM/>^e <http://www.sbc.su.se/~melen/TMHMMfix/>^f Number of transmembrane segments in front of and following the central hydrophilic region in the hydropathy profiles.

enzyme:substrate complex in the absence of the co-ion Na⁺, binding of citrate from the periplasmic or cytoplasmic side of the membrane resulted in different catalytic states with different accessibilities of the thiols in the pore loop. Similarly, binding of the co-ion Na⁺ protected the thiols from the periplasmic side, while the accessibility was only reduced from the cytoplasmic side of the membrane. The data are consistent with a model where the XA region folds in a pore between the transmembrane helices with the cysteine residues accessible from the periplasm for reagents that are small enough to enter the pore. Substrate would bind at a position in the pore in between the cysteine residues and the cytoplasmic side of the membrane (fig. 1). A position of the binding site at the interface of the membrane and cytoplasm is supported by the position of an arginine residue in TMS XI that was demonstrated to interact directly with the bound substrates [Bandell and Lolkema, 2000].

Domain Structure of the 2HCT Transporters

The domain structure in the structural model of the transporters in the 2HCT family is based on sequence analysis of N- and C-terminal parts of the transporters in structural class ST[3]. The family hydropathy profiles of many families in structural class ST[3] are characterized by two hydrophobic regions separated by a central hydrophilic region suggesting that the transporters consist of two membrane-embedded domains connected by a large loop. Moreover, the N-terminal half of one subfamily in the NHAC transporter family in ST[3] showed significant sequence identity with the C-terminal half of another subfamily, suggesting that the two domains might be homologous domains [Lolkema and Slotboom, 2003]. Sequence identity between the N- and C-terminal halves of the proteins in ST[3] was demonstrated by blasting [Altschul et al., 1997] each half sequence against a database containing all the N- and C-terminal half sequences [Lolkema et al., 2005]. Analysis showed that a significant-

ly higher percentage of hits was observed between the N- and C-terminal halves of ST[3] than was observed between both halves and half sequences from unrelated families from structural classes ST[1], ST[2] and ST[4] that were also included in the database. In agreement, sequence homology between the first and second halves of transporters in four families of the IT superfamily, which is part of ST[3], was identified using a different approach [Prakash et al., 2003]. Local alignments between the longest fragments from the N- and C-terminal halves of ST[3] sequences showed that the homology involved the complete membrane-embedded parts. Subsequently, alignment of the averaged hydropathy profiles of these membrane-embedded parts and the family profile of the 2HCT family showed that the two homologous domains comprise TMSs II–VI and TMSs VII–XI. The odd number of TMSs in each domain forces the orientation of the two domains in the membrane to be opposite; the N-terminus of the N-terminal and C-terminal domains resides in the periplasm and cytoplasm, respectively (fig. 1). A second important implication of the two domain structure would be that the pore loop in cytoplasmic loop XA has its counterpart in periplasmic loop VB in between TMSs V and VI in the N-terminal domain. This finding resolved an old issue concerning the membrane topology of CitS of *K. pneumoniae*. Region VB in the sequence of CitS was one of the 12 hydrophobic regions that was predicted to be transmembrane, which, subsequently, was falsified by experiment. Different experimental approaches showed that the hydrophobic region was exported to the periplasm during biogenesis [van Geest and Lolkema, 1996]. Both putative pore loops in the ST[3] transporters are quite hydrophobic, in line with their disposition in between the TMSs. The pore loops contain an extraordinarily high fraction of residues with small side chains (glycine, serine, alanine) which may reflect a compact packing of the loops in between the TMSs. The regions containing the pore loops are among the best conserved regions in the transporter families in ST[3] which emphasizes their role in the transport mechanism.

Conclusion

The combination of experimental work and bioinformatics approaches has resulted in a detailed structural model for the transporters of the 2HCT family. Experiments have demonstrated the presence of 11 *trans* membrane segments, an amphipathic surface helix in between TMSs VIII and IX and a pore-loop structure between

TMSs X and XI. Bioinformatics has identified the domain structure of the proteins and, as a consequence, the pore loop in between TMSs V and VI. How the proposed structure catalyzes translocation of substrates and co-ions is a matter of speculation, but together with a lot of additional experimental data, a mechanism can be hypothesized in which the two pore loops and the adjacent TMSs form the translocation site. The two pore loops would be in close contact in the 3D structure in a single pore that alternately would be opened to either side of the membrane during the catalytic cycle. The substrate-binding site is located at the membrane-cytoplasmic interface, which positions it deep down in the pore when opened to the external face of the membrane. The cytoplasmic pore loop (XA) extends into the pore beyond the binding site, making cysteine residues in the loop accessible from the periplasmic side even when substrate is bound. Opening and closing of the pore to either site of the membrane would be controlled by binding of the substrate and co-ions. The accessibility of cysteine residues in the cytoplasmic pore loop was shown to be different in different catalytic states of the transporter by experiment [Sobczak and Lolkema, 2003, 2004].

References

- Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S 2003. Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* 301:610–615.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search. *Nucleic Acids Res* 25:3389–3402.
- Bandell M, Lolkema JS 2000. Arg-425 of the citrate transporter CitP is responsible for high affinity binding of di- and tricarboxylates. *J Biol Chem* 275:39130–39136.
- Bowie JU 2006. Flip-flopping membrane proteins. *Nat Struct Mol Biol* 13:94–96.
- Cuthbertson JM, Doyle DA, Sansom SP 2005. Transmembrane helix prediction: a comparative evaluation and analysis. *Protein Eng Des Sel* 18:295–308.
- Daley DO, Rapp M, Granseth E, Melén K, Drew D, von Heijne G 2005. Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* 308:1321–1323.
- Dimroth P, Thomer A 1986. Citrate transport in *Klebsiella pneumoniae*. *Biol Chem Hoppe Seyler* 367:813–823.
- Fu D, Libson A, Miercke LJ, Weitzman C, Nollert P, Krucinski J, Stroud RM 2000. Structural determinants of water permeation through aquaporin-1. *Science* 290:481–486.
- Grunewald M, Bendahan A, Kanner BI 1998. Biotinylation of single cysteine mutants of the glutamate transporter GLT-1 from rat brain reveals its unusual topology. *Neuron* 21:623–632.
- Grunewald M, Kanner BI 2000. The accessibility of a novel reentrant loop of the glutamate transporter GLT-1 is restricted by its substrate. *J Biol Chem* 275:9684–9689.
- Huang Y, Lemieux MJ, Song J, Auer M, Wang DN 2003. Structure and mechanism of the glycerol-3-phosphate transporter from *Escherichia coli*. *Science* 301:616–620.
- Hunte C, Screpanti E, Venturi M, Rimon A, Padan E, Michel H 2005. Structure of the Na⁺/H⁺ antiporter and insight into mechanism of action and regulation by pH. *Nature* 435:1197–1202.
- Ishiguro N, Izawa H, Shinagawa M, Shimamoto T, Tsuchiya T 1992. Cloning and nucleotide sequence of the gene (*citC*) encoding a citrate carrier from several *Salmonella* serovars. *J Biol Chem* 267:9559–9564.
- Jack DL, Paulsen IT, Saier MH 2000. The amino acid/polyamine/organocation (APC) superfamily of transporters specific for amino acids, polyamines and organocations. *Microbiology* 146:1797–1814.
- Kästner CN, Schneider K, Dimroth P, Pos KM 2002. Characterization of the citrate/acetate antiporter CitW of *Klebsiella pneumoniae*. *Arch Microbiol* 177:500–506.
- Kawai S, Suzuki H, Yamamoto K, Kumagai H 1997. Characterization of the L-malate permease gene (*maeP*) of *Streptococcus bovis* ATCC 15352. *J Bacteriol* 179:4056–4060.
- Khademi S, O'Connell J III, Remis J, Robles-Colmenares Y, Miercke LJW, Stroud R 2004. Mechanism of ammonia transport by Amt/MEP/Rh: structure of AmtB at 1.35 Å. *Science* 305:1587–1594.
- Krogh A, Larsson B, von Heijne G, Sonnhammer EL 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* 305:567–580.
- Krom BP, Aardema R, Lolkema JS 2001. *Bacillus subtilis* YxkJ is a secondary transporter of the 2-hydroxycarboxylate transporter family that transports L-malate and citrate. *J Bacteriol* 183:5862–5869.
- Krom BP, Lolkema JS 2003. Conserved residues R420 and Q428 in a cytoplasmic loop of the citrate/malate transporter CimH of *Bacillus subtilis* are accessible from the external face of the membrane. *Biochemistry* 42:467–74.
- Lolkema JS, Slotboom DJ 1998. Estimation of structural similarity of membrane proteins by hydropathy profile alignment. *Mol Membr Biol* 15:33–42.
- Lolkema JS, Slotboom DJ 2003. Classification of 29 families of secondary transport proteins into a single structural class using hydropathy profile analysis. *J Mol Biol* 327:901–909.
- Lolkema JS, Slotboom DJ 2005. Sequence and hydropathy profile analysis of two classes of secondary transporters. *Mol Membr Biol* 22:177–189.
- Lolkema JS, Sobczak I, Slotboom DJ 2005. Secondary transporters of the 2HCT family contain two homologous domains with inverted membrane topology and *trans* re-entrant loops. *FEBS J* 272:2334–2344.
- Marty-Teyssset C, Lolkema JS, Schmitt P, Divies C, Konings WN 1995. Membrane potential-generating transport of citrate and malate catalyzed by CitP of *Leuconostoc mesenteroides*. *J Biol Chem* 270:25370–25376.
- Molenaar D, Bosscher JS, ten Brink B, Driessen AJ, Konings WN 1993. Generation of a proton motive force by histidine decarboxylation and electrogenic histidine/histamine antiporter in *Lactobacillus buchneri*. *J Bacteriol* 175:2864–2870.
- Murata K, Mitsuoka K, Hirai T, Walz T, Agre P, Heymann JB, Engels A, Fujiyoshi Y 2000. Structural determinants of water permeation through aquaporin-1. *Nature* 407:599–605.
- Pornillos O, Chang G 2006. Inverted repeat domains in membrane proteins. *FEBS Lett* 580:358–362.
- Pornillos O, Chen YJ, Chen AP, Chang G 2005. X-ray structure of the EmrE multidrug transporter in complex with a substrate. *Science* 310:1950–1953.
- Prakash S, Cooper G, Singhi S, Saier MH Jr 2003. The ion transporter superfamily. *Biochim Biophys Acta* 1618:79–92.
- Rapp M, Granseth E, Seppälä S, von Heijne G 2006. Identification and evolution of dual-topology membrane proteins. *Nat Struct Mol Biol* 13:112–116.
- Saier MH Jr 2000. A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* 64:354–411.
- Saier MH Jr, Beatty JT, Goffeau A, Harley KT, Heijne WH, Huang SC, Jack DL, Jahn PS, Lew K, Liu J, Pao SS, Paulsen IT, Tseng TT, Virk PS 1999. The major facilitator superfamily. *J Mol Microbiol Biotechnol* 1:257–279.
- Slotboom DJ, Konings WN, Lolkema JS 1999a. Structural features of the glutamate transporter family. *Microbiol Mol Biol Rev* 63:293–307.
- Slotboom DJ, Sobczak I, Konings WN, Lolkema JS 1999b. A conserved serine-rich stretch in the glutamate transporter family forms a substrate-sensitive reentrant loop. *Proc Natl Acad Sci USA* 96:14282–14287.
- Sobczak I, Lolkema JS 2003. Accessibility of cysteine residues in a cytoplasmic loop of CitS of *Klebsiella pneumoniae* is controlled by the catalytic state of the transporter. *Biochemistry* 42:9789–9796.
- Sobczak I, Lolkema JS 2004. Alternate access and a pore-loop structure in the Na⁺-citrate transporter CitS of *Klebsiella pneumoniae*. *J Biol Chem* 279:31113–31120.
- Sobczak I, Lolkema JS 2005a. The 2-hydroxycarboxylate transporter family: physiology, structure, and mechanism. *Microbiol Mol Biol Rev* 69:665–695.
- Sobczak I, Lolkema JS 2005b. Structural and mechanistic diversity of secondary transporters. *Curr Opin Microbiol* 8:161–167.
- Van Geest M, Lolkema JS 1996. Membrane topology of the Na⁺-dependent citrate transporter of *Klebsiella pneumoniae*. Evidence for a new structural class of secondary transporters. *J Biol Chem* 271:25582–25589.
- Von Heijne G, Lolkema JS 1999. Insertion of a bacterial secondary transport protein in the ER membrane. *J Biol Chem* 274:2816–2823.
- Wei Y, Guffanti A, Ito M, Krulwich TA 2000. *Bacillus subtilis* YqkI is a novel malic/Na⁺-lactate antiporter that enhances growth on malate at low protonmotive force. *J Biol Chem* 275:30287–30292.
- Yamashita A, Singh SK, Kawate T, Jin Y, Gouaux E 2005. Crystal structure of a bacterial homologue of Na⁺/Cl[−]-dependent neurotransmitter transporters. *Nature* 437:215–223.
- Yernool D, Boudker O, Jin Y, Gouaux E 2004. Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* 431:811–818.